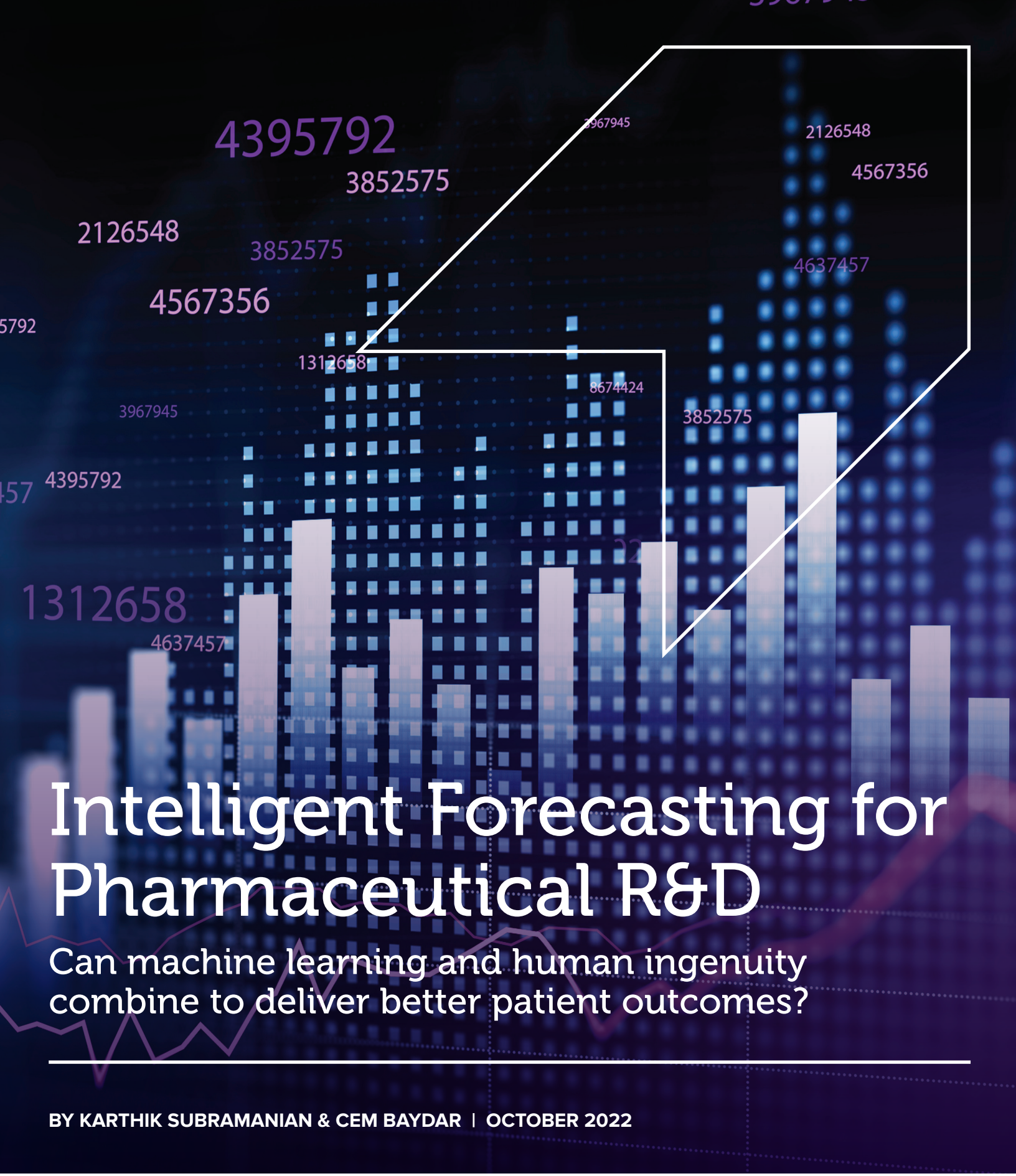

Intelligent Forecasting for Pharmaceutical R&D

Can machine learning and human ingenuity combine to deliver better patient outcomes?

BY KARTHIK SUBRAMANIAN & CEM BAYDAR | OCTOBER 2022

Evaluate
a norstella company

Contents

Executive Summary

Commercial analytics and forecasting underpin strategic decision-making and business planning across the pharmaceutical and life science sectors. Inaccurate risk assessments and poor strategic decision making cost the industry billions of dollars in wasted research and development (R&D) spend every year and result in a gap in pharmaceutical intelligence to inform decision making in the clinical development space. What's needed is the ability to forecast the commercial potential of early-stage assets that;

- Reduces the need for complex manual forecasting
- Allows products to be valued at scale
- Delivers the speed and accuracy needed to make time-sensitive licensing or R&D portfolio decisions with confidence
- Provides a good understanding of probabilities of technical and regulatory success for those assets that are in clinical development

Intelligent forecasting, unlike traditional forecasting methodologies, considers the past, present and future drivers of commercial and clinical success in real time. It is a forecasting methodology powered by machine learning and predictive analytics that allows complex real-world clinical development and commercial questions to be addressed dynamically and with up-to-date information.

Intelligent forecasting considers the important aspects of the pharmaceutical value chain such as;

- Complexities of clinical development
- Non-linear relationships between product attributes
- Randomness and uncertainties of drug discovery and market dynamics
- Companies' current development pipeline content and strategic fluctuations over time.

Intelligent forecasting applies machine learning technologies to lever millions of historical clinical and commercial events and this innovative approach has the ability to push the boundaries on pharmaceutical intelligence generation. Evaluate Omnium uses machine learning to provide real-time and granular product/indication-level predictions for early- and mid- stage development assets.

Peak sales predictions can be accurately generated using intelligent forecasting, whether a development asset is in Phase I, II or III. A detailed performance analysis of an intelligent forecasting platform yielded an accuracy rate of 88%, i.e. almost nine out of 10 predicted peak sales fall within a tolerated range acceptance of the actual peak sales achieved[1]. These predictions are made when a product is in clinical development, and as such these sales estimates are made 12 to 15 years in advance of actual peak sales occurring.

Probability of technical and regulatory success predictions, i.e., clinical development success rates, can be analysed at a product-level based on therapeutic indication and are available for Phase I, II or III milestones at the start of each phase as well after phase transitions. There is circa 65-77% prediction accuracy across all phases with estimates indicative of intelligent forecasting producing accurate predictions for c. three out of four events i.e., trial success or failures.

There is a further increase in prediction accuracy when probabilities are analysed at therapeutic area level, including for those that are of current strategical importance for pharmaceutical companies. For example, in oncology for phase II we observe 84% accuracy in predicting trial failure and 78% accuracy in predicting trial success. Conventionally used benchmarks for success rates do not include events from the last 7 years whereas Evaluate Omnium draws upon both historical and real time events to ensure relevance to current pharmaceutical companies' development pipelines.

> Intelligent forecasting can provide real world decision support that customers can use to inform decision making across the entire life cycle of a therapeutic product, including preclinical discovery, overcoming historical and emerging forecasting challenges.

1. 88% of predicted peak sales fall within an acceptable range of the actual US peak sales achieved (sales value +/- one median standard deviation)

Forecasting and insight generation is foundational to decision making across the life science and healthcare ecosystem

Commercial analytics and forecasting underpin strategic decision-making and business planning across the pharmaceutical and life science sectors. Pharmaceutical companies and their supply chain, financial analysts and consultancies from across the sector make significant investments seeking to secure access to accurate forecasting data, as well as robust and relevant insights into their clinical success rates. These data and insights are essential for managing and growing their businesses and making informed decisions around future investments.

Drug development is a lengthy, risky and expensive process. Even if you guarantee technical and regulatory success, commercial success is far from certain and the opportunity to make returns is not limitless. Therefore, good decisions must be taken to ensure successful portfolios. There is market demand for relevant insights into commercial opportunities and a need to gain a good understanding of probabilities of technical and regulatory success for pharmaceutical assets that are in clinical development (e.g., Phases I, II and III).

With a patent cliff in 2028 looming, the race is on to identify drugs that will drive the next wave of success. Just 20% of marketed drugs generate 90% of commercial returns, suggesting that pharmaceutical companies' investment in novel molecules may not always deliver the returns expected[2]. With average peak sales of a blockbuster drug reaching $2.8 billion, there are opportunities to add tens of billions in commercial value if decision makers can accurately forecast how likely a drug will be to reach market and how long it will take to get there, while identifying its true commercial return potential.

Unfortunately, the failure to comprehensively assess the potential increases the risk of forecasting errors – with a huge impact on revenue. Evaluate's[3] study into the accuracy of forecasting found that:

1. Every 1% error in under-forecasting results in an estimated $200 million of lost sales revenue = $69 million EBIT and $54 million of net income.
2. Every 1% error in over-forecasting results in an estimated $93 million of additional cost commitments ($46 million in cost of goods/COGs and $47m in selling, general and administrative expenses/SG&A) across the portfolio.

To proceed with confidence in this daunting environment, companies need to have a complete and accurate understanding of the drugs in the pipeline.

2. The patent winter is coming | Evaluate
3. The patent winter is coming | Evaluate

How does the early-stage information gap limit accurate commercial forecasting?

Initiating equity analyst coverage in late-stage development makes sense; there is little point in creating detailed forecast models for products that may never reach the market, or that are not likely to contribute to company revenues within the forecast horizon. But this tendency also means that pharmaceutical companies are often without independent, readily available valuation data to inform their pipeline decisions.

With the bulk of R&D spending required in Phase III, plus the higher price tags that later-stage assets command in licensing or M&A transactions, earlier insight into product value could dramatically increase the productivity of R&D investments.

Third-party valuations from equity researchers or consensus forecast providers have long been the go-to option for pharmaceutical companies. In many cases, this is fine, but it does tend to leave a gap around coverage for early-stage assets. A recent analysis of equity analyst forecast availability showed that of blockbuster product launched in the past decade, 56% lack forecast prior to Phase III. For non-blockbuster products, only 39% had coverage even at Phase III – see Figure 1.

Figure 1: Big Pharma Product Coverage by Phase
Drugs launched between 2010–2021

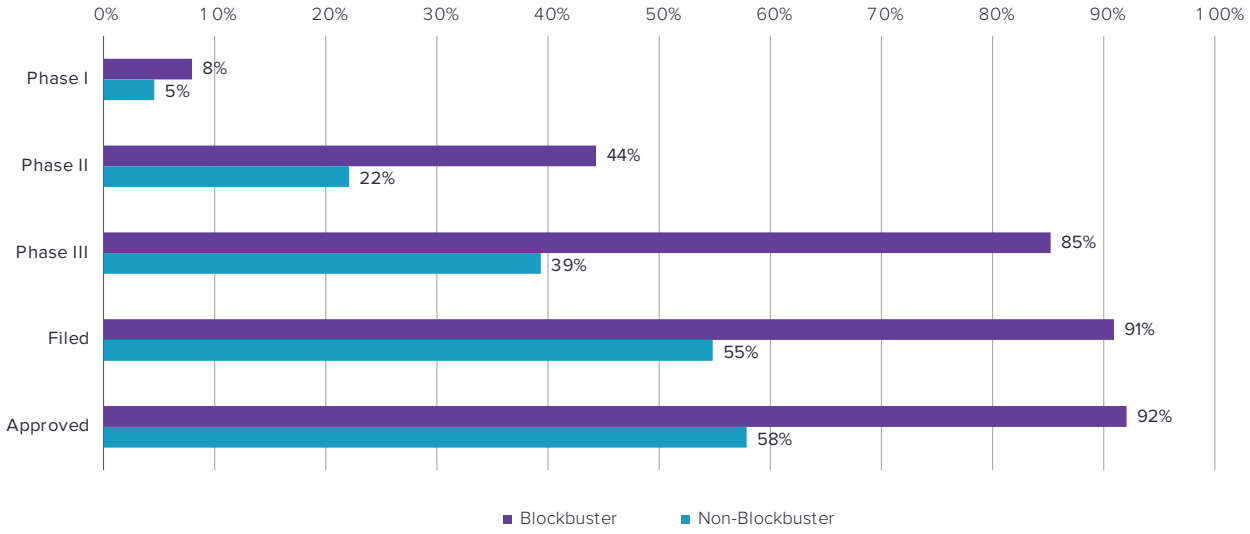

Figure 1: recent Evaluate analysis of equity analyst forecast availability across the pharmaceutical value chain, covering 282 Blockbuster drugs and 704 Non-Blockbuster drugs

What are the drivers to inform better decision making for products in clinical development?

Pharmaceutical companies require product and indication forecasts and predictions that are evidence based. These forecasts need to incorporate context such as historical and recent events whilst taking into account both clinical and commercial activities so they can accurately predict the future clinical and commercial outcomes.

Based on our experience for assets in clinical development, there are three key areas that require evidence-based insights to inform strategic decision making:

1. **Predictors of technical and regulatory success i.e., clinical development success rates** - Pharmaceutical and biotechnology companies and analyst organisations require:
   i. Predictions of technical and regulatory success for development asset transitions from Phase I-to-Phase II, Phase II-to-Phase III and Phase III-to-Filing. These estimates of clinical development success need to be at the indication level
   ii. Clinical trial durations for each phase of clinical development at an asset and importantly the disease indication level
   iii. Transition times between phases needs to consider clinical trial design, therapeutic area and asset modality

   When you take these requirements together these predictions provide estimates for development time and probabilities of success.

   Organisations currently use estimates of success based on their historical in-house development programs, if they are a pharmaceutical company, or alternatively from published benchmarks estimated from historical clinical development success rates. It should be noted that these public analyses are generally based on historical datasets and are relatively high level when it comes to indication e.g., oncology or cardiovascular rather than more granular insights on indications such as non-small cell lung cancer (NSCLC) or coronary artery disease (CAD). Some of the in-house product development benchmarks are also biased in many ways.

2. **Commercial opportunity forecasts for early- to late-stage clinical development assets** – Pharmaceutical and biotechnology companies and analyst organisations all require estimates of commercial opportunities before a product is approved and launched, often even before the product label is in development. These estimates can be used to inform R&D investments, portfolio prioritisation, and to understand commercial opportunities associated with developing target product profiles (TPPs). They can also be used to aid product, portfolio or company valuations calculated using predictors of technical and regulatory success to assess risk considering estimated peak sales.

   Commercial estimates currently use insights gained from a variety of sources including volume- or epidemiology-based forecasts or consensus estimates. There are few sources of commercial opportunity forecasts for early- to late-stage clinical development assets unless bespoke forecast has been developed, and users generally are relying on published market research reports that describe sales opportunities. These reports are at top level therapeutic areas and with consensus estimates the authors have applied risk adjustments that may not be apparent to users of the commercial estimates (See Table 2 for a discussion of the advantages for each approach). These conventional, and often complex, manual forecasting methodologies do not allow products to be valued at scale.

3. **Probability of technical success, i.e., preclinical to clinical transition rates, and commercial insights into risks and commercial opportunities for preclinical candidates or disease indications** – Pharmaceutical and biotechnology companies need to understand early-stage attrition and inform early-stage strategic decision making. As such they need to understand and then build confidence in rationale, in mechanistic safety whist gaining an understanding of commercial opportunities, picking the targets and candidates that will survive.

   These types of probability insights are bespoke, based on a company's historical attrition rates or published success rate benchmarks and often the experience of their teams. Commercial opportunity insights are often gained from a mix of primary market research, disease registries and published analyst reports.

Do the conventional forecasting methodologies address the complexity challenges and dynamic nature of clinical development?

Pivotal to forecasting success is a consistent high quality historical data, that is supplemented with real time clinical and commercial events. Valuable insights into commercial value and probabilities of success can be estimated when this foundational dataset is overlayed with robust forecasting methodologies that can identify signals of potential commercial and clinical success for early, mid, and late-stage pipeline products.

The forecasting model itself needs to have broad applicability but must also be able to deep dive into disease indication to provide outputs at the product-, indication- and modality-levels. This will ensure that the outputs address the specific needs of the life science sector whether it is a pharmaceutical or biotech company, analysts/investors, or consultancies. Users of these models need to understand the assumptions on how the forecasting is derived as building confidence in the outputs is paramount if strategic decision are to be made on these forecasts.

The forecasting model needs to be flexible enough to consider new parameters and be able to refine existing parameters to improve the model predictability, to reflect the dynamic nature of drug development and maintain output relevance. This is pertinent as there is a strategic shift from small molecule medicines (e.g., oral, tablets) that are prescribed in the primary care setting toward one of prescribing complex molecules (e.g., biological, injections) for specialty care indications. More recently there has been a further shift from symptom relief to disease transformation and even potentially curative approaches. Forecasting models need the ability to accurately factor in impact of future changes in the wider market, such as new competition, regulatory changes, etc.

As these are predictions used to inform strategic decision making, it is essential that there is a good correlation between prediction and reality e.g., predicted peak sales should reflect actual peak sales reported (although pricing and differences between net and list prices may impact prediction accuracy) or probabilities of technical success verses actual success rates for a novel mechanism of action development candidate.

Whilst the conventional bottom-up forecasting methodologies such as volume-, epidemiology- or consensus-based forecasting are very good at providing a commercial or financial model, they are largely inadequate in modelling complex forecasting processes that consider the past, present and future drivers of commercial and clinical success.

When you consider the complexity and dynamic nature of drug discovery it becomes apparent that the conventional methodologies to forecasting will have limitations in contrast to intelligent forecasting and machine learning approaches which are far more dynamic and adaptable to real time events and trends. The differences are illustrated in the Table 1 below:

Table 1

Conventional (bottom up) forecasting e.g., volume-, epidemiology- or consensus-based forecasting	**Intelligent forecasting** e.g., machine learning and predictive analytics
Conventional (bottom up) forecasting methodologies, generally using spreadsheets, are utilised to provide a heavily simplified version of the real-world clinical development and market commercial opportunities and this process is: • Static and required actual values • Linear progression from phase I to II, II to III and III to filing • Deterministic - no randomness is considered in the forecast • Independent as not seen as depending on other variables • Abstract – theoretical • Complex and manual	Intelligent forecasting is a new modelling solution, that deploys predictive analytics and machine learning in conjunction with real time big data methodologies, that allows the development of complex real-world problems in a digital world, considering important aspects of the physical world such as: • Complex feedback loops between timelines, probabilities, and commercial opportunities • Non-linear relationships of different entities, modalities, and stages of development • Randomness and uncertainties associated with drug discovery and development • Market dynamics – self-adaptive models learning and updating from new evidence • Other unknown future states

Table 1: Comparison of conventional verses intelligent forecasting in terms of addressing the complexity challenges and dynamic nature of drug discovery - Intelligent forecasting will provide real world decision support that customers can use to inform decision making across the entire life cycle of a therapeutic product, including preclinical discovery, overcoming historical and emerging forecasting challenges

Table 2 shows a description of forecasting methodologies and summarise the strengths, weakness, opportunities, and threats to each approach.

Table 2

	Volume-based forecasting	Epidemiology-based forecasting	Consensus forecasting	Intelligent forecasting
Description	• Based on Rx or prescription volumes	• Patient based approach relying on layers of assumptions to generate a forecast	• Aggregating broker forecasts (usually patient based) to provide a consensus view of the market	• Analysing historical datasets to identify signals of potential success in pipeline products, translated into a dollar value
Strengths	• Important for tactical forecasting purposes (e.g., manufacturing, sales estimates, calibration of other forecasts) • Simple and easy to understand	• Important mainstay for both tactical (launch planning & patent expiry) and strategic purposes (market landscaping & sizing) through all phases of the lifecycle • Can follow the layering of assumption s • Long range (6+ years) • Considered the gold standard in forecasting	• Mid-Long range (4-7 years) • Tend to become significantly less accurate over the long run • Strategic uses landscaping for market sizing & pipeline evaluations • Useful for market landscaping	• Significantly enhanced pipeline coverage compared with other methodologies (Phase I – Filed) • A balanced and fair approach to each individual product (ideally should reduce bias) • Timely reactions and updates in response to news flow • Provides a new data point which cannot be sourced internally for triangulation purposes
Weaknesses	• Short range (~18 months) • Tend to become significantly less accurate over the long run • Only available for launched products with a sales history	• Tend to be driven by more complex assumptions and data requirements (varies by model) • Syndicated forecasts are event driven and can only cover late-stage products (Phase III & Filed)	• Limited insights into the drivers and assumptions underpinning the models • Brokers can interpret data differently resulting in divergent results • Mostly limited to Phase III & Filed with some coverage of early stage	• A single error can be compounded throughout the whole model • Model can struggle to handle new situations if not constantly reviewed and trained (e.g., new technologies, evolving disease areas)
Opportunities	• Can be used to great effect in more mature product markets (e.g. post launch) • Important for predicting seasonality	• Can be highly effective if good data is available • Extremely effective at modellling specific events	• Access to underlying assumptions such as price, share and patients would significantly improve the value of the offering	• Extensive opportunities for innovation to cover more products and include more features • Can be scaled to other markets more easily
Threats	• Significantly less accurate	• Limited scalability due to data requirements and nuances across markets	• Forecasts are all risk adjusted which limits the usefulness of R&D forecasts • Coverage is unlikely to increase, adding more brokers is likely a case of diminishing returns	• Acceptance is more of a challenge • Potential distrust of the “black box” approach despite transparency

Table 2: A comparison of the different forecasting methodologies (SWOT analyses)

What makes Intelligent Forecasting highly reliable and how does it address the "black-box" myth?

Intelligent forecasting does not need to be a "black box" – it can be a transparent methodology that harnesses the predictive power of machine learning to generate an accurate and comprehensive outputs that inform strategic decision making.

Near real-time data, coupled with a flexible and dynamic approach to forecasting means the outputs are current and relevant to pharmaceutical companies' current strategic direction e.g., speciality care, multi-indication products, complex modalities and precision medicine. Intelligent forecasting consists of seven individual but linked processes. These are illustrated in Figure 2 and summarised below.

1. **Curation of foundational dataset from high confidence sources** Granular historical pipeline coverage including individual trial, phase transition timelines and attritions, key R&D metrics and historical sales, product characteristics, product news flows and results, company characteristics, competitive environment, market news flows and results.

2. **Pre-forecast data processing** – this includes standardisation and structuring in order that the data is suitable for analyses. Events are assigned to the model timeline and coded, then datasets extracted to first create training datasets and then the data that feeds the machine learning ahead of the daily running of the intelligent forecasting.

3. **Development on training and test datasets** – a balanced and unbiased subset of the data that is used to train the machine learning and form the basis of estimating the predictability of each attribute. The test dataset is used to validate the accuracy of the machine learning model

4. **Defining product and indication attributes** – these are characteristics that define the product and indication selected for forecasting. There are four categories of attributes –
   i. Product characteristics e.g., modality sales potential, clinical trial milestones (to date), etc.
   ii. Company characteristics e.g., track record and historical success rates, therapeutic area focus, etc.
   iii. Unmet need within the indication, number of mechanisms of action, orphan drug status, fast track designation, ratio of approvals to failures in indication, etc.
   iv. Competition within the indication e.g., number of products in development or launched, number of different mechanisms of action, etc.

   Taken together these attributes define the development probabilities of success and commercial potential of a product in development.

5. **Attribute selection based on statistical analysis of predictability and attribution correlations** – this predictability of model outputs based on the training dataset, and the correlations to real world events. It defines which attributes should be included in the machine learning model and can be flexed depending on product type and indication. The statistical analysis allows for the selection of a series of 50 to 150 predictive attributes for inclusion in the model. The analysis firstly informs which attributes can be considered statistically relevant, and secondly sets the predictive accuracy for a model. It is important that the predictability and correlation of these attributes are regularly reviewed to reflect changes in pharmaceutical strategy.

6. **Intelligent forecasting methodology applied to real time clinical and commercial events to predict product/indication level outputs** - where the machine learning methodologies feed real-time data to generate product/indication-level probabilities of technical and regulatory success and/or forecasts of commercial opportunity. For peak sales predictions the model uses a regularised regression model with a log transformed target variable. A supervised learning classification algorithm is instead used to predict the clinical trial phase progression probabilities of technical and regulatory success.

7. **Forecast outputs are transparently** produced, quality assured and tested against industry benchmarks to ensure consistency. An audit of drivers can also be performed where it is possible to identify which attributes are used to predict success rates or commercial opportunities. The impact of each attribute on the prediction can be understood though the use of 'leave one out' or 'one at a time' approaches.

8. **Forecast outputs are flexible that can be viewed in various ways** – Macro-level Technology, Disease or Company landscape predictions to micro-level granular product-indication predictions.

Evaluate Ltd has built up significant historical data and years of sophisticated R&D effort to come up with a new way of intelligently forecasting of pipeline assets. Evaluate Omnium is a new way to assess pharmaceutical asset potential by giving users an unprecedented view of both the predicted risk (likelihood of success) and return (commercial potential) of pharmaceutical assets in all phases of clinical development. Our analysis of the results provided an unparalleled insight into drug portfolio potential at a product-specific level, including early-stage and privately-owned drugs.

Evaluate Omnium uses machine learning to analyse historical datasets to identify signals of clinical success for products at all stages of the pipeline. This paper looks at how these probabilities, when combined with historical and current commercial datapoints, are capable of delivering commercially valuable insights into sales estimates for development assets often 12 to 15 years in advance of actual peak sales occurring.

Figure 2

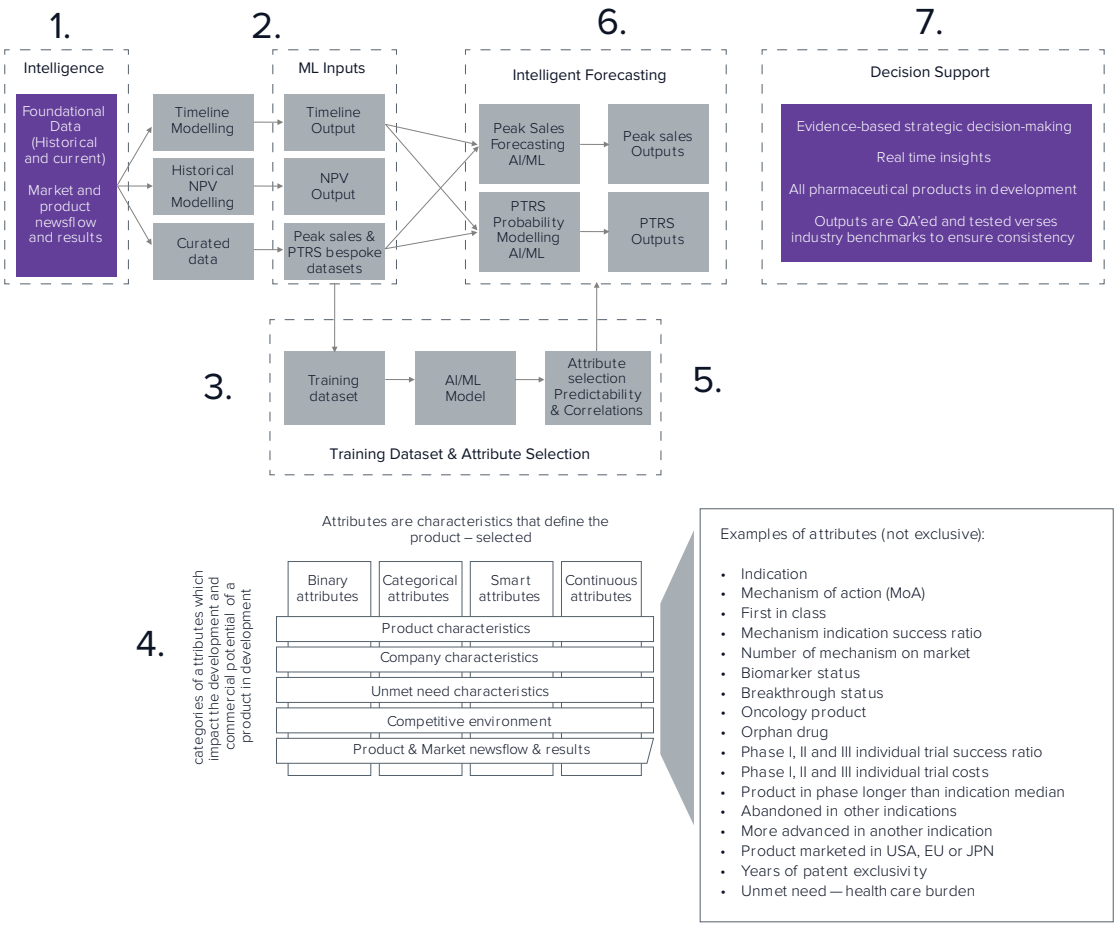

Figure 2: Intelligent forecasting using Evaluates Omnium – End to end process from data harvesting and curation, through to machine learning-based forecasting and generation of evidence-based insights on which decision scan be confidently made

How do pharmaceutical and life science organisations use intelligent forecasting?

1. **A prominent biotechnology organisation required a specific-purpose analytical model to predict the probability of technical success for over 100 early phase assets they are evaluating to acquire.** The advantage of using intelligent forecasting in this situation was that the generalised models, and benchmarking approaches, mostly undervalue the different nuances that built-up based on aggregated data. Widely used models do not enable the organisations assessment team to understand the different driver parameters such as orphan drug designation, technology and mechanism of action which makes them mostly a black-box model. This intelligent forecasting approach also enabled the organisation to both validate their assumptions as well as identify differences and the driving reasons in these situations.

2. **A mid-size pharmaceutical company was interested in understanding the potential peak sales and expected net present value (NPV) of their key early phase asset in their portfolio.** Intelligent forecasting was used to forecast different potential uptake dynamics under different scenarios. Intelligent forecasting was supplemented with a bottom-up epidemiology-based model to assess different commercial return options and enabled them to prioritise different indication options.

3. **A hedge fund that sees an opportunity in public biotech companies trading at a discount.** They have identified a set of around 30 companies that they would like to analyse further before taking any positions. As generalists with some healthcare experience rather than specialist healthcare investors, they're relying on Evaluate Omnium to provide them with comprehensive data backed by rigorous methodologies and domain expertise. They are using PS-PTRS and timelines as a data point to help them determine which drugs/companies are likely to succeed fail.

4 **A pharmaceutical company, with a product that was approaching patent cover expiry, needed to understand the potential impact of loss of exclusivity.** It is not only the early-stage assets where you need to apply intelligent forecasting algorithms. The company had a key product approaching patent expiry and they need to assess the value at-risk in order to estimate the potential value erosion. Intelligent forecasting was used to identify analogue products that have gone through the same market conditions such as competitive intensity and unmet needs. This helped the company forecast the potential outcomes and estimate the commercial impact, while also assessing several options to mitigate the risk.

How robust is intelligent forecasting and how good at predicting clinical and commercial events is a machine learning approach?

One of the big four firms conducted a series of performance testing analyses on the machine learning methodologies that underpin Evaluate Omnium. The goal was to understand how robust the predictions and forecasts generated compare to actual events. These illustrate the value of intelligent forecasting to organisations decision making and the accuracy of clinical and commercial predictions.

Evaluate Omnium is capable of generating peak sales forecasts with an accuracy rate of c.87%, for development assets across the clinical development value chain. These forecasts are produced at the individual product-level, are dependent on therapeutic indication and consider recent market attributes.

Evaluate Omnium generates product-level peak sales predictions that are produced whether a development asset is in Phase I, II or III. In addition, these peak sales forecasts can be estimated at more granular disease-level indications e.g., solid tumours or blood cancers rather than broad cancer.

Peak sales estimates consider all aspects of the pharmaceutical value chain including the non-linear relationships between product attributes, the often

Figure 3: Intelligent forecasting for pharmaceutical research and development

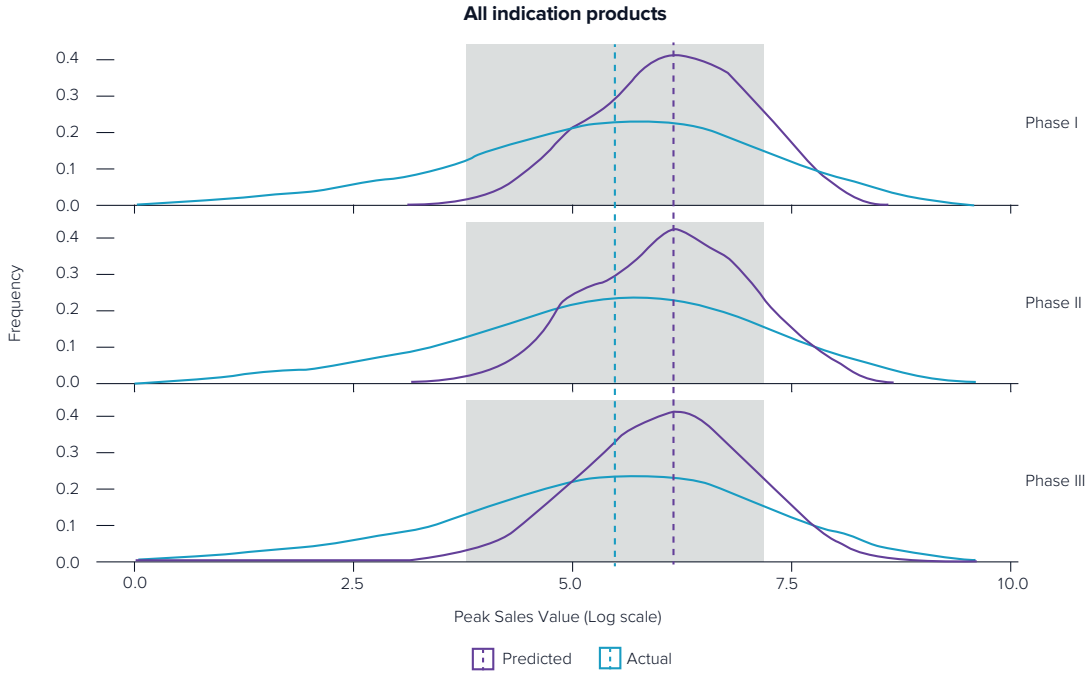

Figure 3: Distribution of target versus predicted values across phases for all indication products. The dotted lines correspond to the observed median values and the vertical grey shaded area represents the area within one median standard deviation of the predicted median. Values are displayed in log scale. Sample size was 1600 products

randomness and uncertainties of drug discovery and market dynamics, as well as companies' current development pipeline content and strategic fluctuations over time.

Performance analysis shows an accuracy rate of circa 87-88% for predicted peak sales, for products in Phase I, II and III (analyses based on 1600 products). This accuracy rate is based on the percentage of products where predicted peak sales fall within an acceptable range of the actual US peak sales achieved (sales value +/- one median standard deviation). These predictions are made when a product is in clinical development, and as such these sales estimates are made 12-15 years in advance of actual peak sales occurring.

Looking at the performance of the forecast from Evaluate Omnium, there is an overlapping distribution of predicted and actual peak sales - see Figure 3. It is apparent that the actual peak sales values (purple curve) are characterised by higher level of dispersion around the median (represented by the dotted lines), indicating that the actual sales are more broadly distributed respect to the predicted values (blue curve) and characterised by lower amounts. The products at the high end (right tail of distribution) are likely to included products with multi-indications, whereby an indication by indication increase in sales is built over time. In contrast, products within the lower end of the distribution (left tail of distribution) are likely to include products with unexpectedly low sales or where marketing support has been withdrawn. The predicted median value (grey dotted line) is generally above the target median across all phases (orange dotted line), suggesting that there is a tendency of intelligent forecasting to generate higher peak sales across all phases. However, circa 87% of predicted peak sales, for development products in Phase I, II and III, fall within an acceptable range of the actual peak sales median value (+/- one median standard deviation[4] represented by the vertical orange shaded area in Figure 3).

In another analysis, comparing with consensus forecasts on early-stage vs late-stage assets, Evaluate Omnium has shown 53% better accuracy than consensus to forecast the commercial potential of assets in the early stages of the clinical pipeline (see Figure 4). Further analyses showed that intelligent forecasting is largely comparable, or slightly less accurate (-8%) than consensus when forecasting commercial potential for late-stage pipeline assets. It should be noted that consensus coverage rapidly diminishes as one moves further back in to the mid- and then early-stage pipeline and that intelligent forecasting delivers a more accurate view on asset potential for early-stage assets (Phase I & II). Consensus continues to play a major role for both marketed as well as close to market assets and continues to have significant strategic value when assessing growth and direction of the overall market.

Figure 4: Analysis of early-stage vs late-stage products

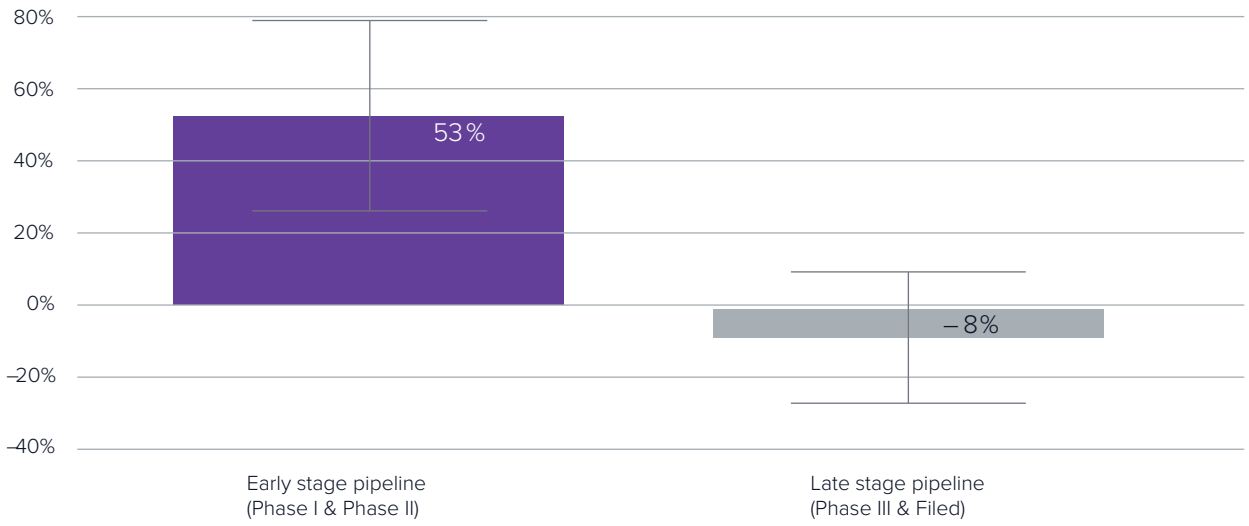

Figure 4 - Analysis based on comparison of Mean Square Error (MSE) for Intelligent forecasting and historical analysis of Consensus Forecasts compared to actual product sales over past 10 years. Analysis of early-stage products are shown in left hand column and late stage products in right hand column

4. The median standard deviation differs from the common standard deviation in that the mean in replaced by the median, which is showed to be more resilient than the mean to the presence of outlying observations

Evaluate Omnium generates clinical development success rates that are indicative of probability of technical and regulatory success. These predictions are produced at the individual product-level, are dependent on therapeutic indication and consider recent clinical development events

Probability of technical and regulatory success predictions can also be generated at product-level based on therapeutic indication. These product-level predictions are produced for Phase I, II or III milestones at the start of each phase as well after phase transitions.

Pharmaceutical companies, business analysts and consultancies generally rely on, and widely cite, benchmark publications of probability of success to inform their decision making. Intelligent forecasting using Evaluate's Omnium can generate probability of technical and regulatory success predictions that are broadly in line with other publications and available benchmarks (e.g., Evaluate PTRS benchmarks[5], Wong et al, 2019[6], BIO/BioMedTracker/Amplion ('BIO 2016')[7]). For Evaluate's Omnium platform we estimated the confidence intervals around the median point of the predicted probability distribution and then to compare these to publications by BIO 2016 and Wong et al, 2019 (see Table 3). Evaluate's Omnium predictions are provided as a range because of there being multiple indications per therapeutic area and that intelligent forecasting is capable to estimating clinical development success rates at these indication-level. As such, intelligent forecasting produces a range of probabilities when we aggregate back up to the therapeutic area level.

Table 3

Therapeutic Area Name	Phase I to II				Phase II to III				Phase III to Filing			
	Omnium Estimates	Evaluate Benchmarks	BIO 2016	Wong 2019	Omnium Estimates	Evaluate Benchmarks	BIO 2016	Wong 2019	Omnium Estimates	Evaluate Benchmarks	BIO 2016	Wong 2019
Cardiovascular	67.4 – 74.6	71	60	73	34.3 – 40.0	52	24	66	34.8 – 52.3	59	47	62
Immunology	64.3 – 74.0	63	66	70	31.4 – 40.0	48	32	46	54.9 – 73.8	56	54	64
Infections	75.0 – 80.0	73	70	70	46.1 – 52.4	60	43	58	61.4 – 68.4	66	65	75
Oncology	76.0 – 77.5	70	63	58	20.5 – 22.5	31	25	33	45.8 – 61.7	56	33	36
Psychiatry	63.3 – 70.5	65	54	n/a	17.6 – 23.8	39	24	n/a	25.4 – 39.1	54	49	n/a
Respiratory	60.1 – 68.0	67	65	n/a	16.6 – 23.5	40	29	n/a	40.7 – 71.6	62	67	n/a

Table 3: Probability of technical and regulatory success estimates produced by Evaluate Omnium Estimates of clinical development success rates with the confidence intervals around the median point of the predicted probability distribution for intelligent forecasting verses widely cited benchmark publications from by BIO 2016 and Wong et al, 2019.

The Evaluate Omnium data is by far the largest following over 17 years of data collection covering c.1660 products, 14,620 phase I transitions, 13,040 phase II transitions and 2,950 phase III transitions.

5. Evaluate Omnium PTRS Benchmarks
6. Wong et al (2019), Estimation of clinical trial success rates and related parameters Biostatistics Apr 1;20(2):273-286 DOI: 10.1093/biostatistics/kxx069
7. Clinical Development Success Rates 2006-2015 - BIO, Biomedtracker, Amplion 2016.pdf

Evaluate Omnium estimated clinical development success rates are indicative of accurately predicting c.3 out of 4 events in clinical development i.e., trial success or failures.

Performance analysis of Evaluate Omnium intelligent forecasting shows that there is c.65-77% prediction accuracy across all phases of clinical development based on analyses of 14,620, 13,040 and 2,950 events for Phase I, II and III respectively. These estimates are indicative of intelligent forecasting producing accurate predictions for c.3 out of 4 events in clinical development i.e., trial success or failures.

Evaluate Omnium can be accurate under different test scenarios, especially in Phase II and Phase III when we consider clinical trial outcomes in the entirety, where predictions continue to surpass 70% in accuracy:

- Phase I: 65.3%
- Phase II: 67.8% - 75.5%
- Phase III: 72.4% - 77.3%

Clinical success rates increase further, to c.88% accuracy, when estimates are considered at a therapeutic area level.

Of particular interest is the increase in prediction accuracy when probabilities of predicting success, or predicting failure, are analysed at the individual disease indications for therapeutic areas that are of current strategical importance for pharmaceutical companies. For example, oncology where we observe higher accuracy predictions at the therapeutic area level than we do when we compared all clinical trial outputs:

- Phase I - 64% accuracy in predicting oncology trial failure and 79% accuracy in predicting trial success
- Phase II - 84% accuracy in predicting oncology trial failure and 78% accuracy in predicting trial success
- Phase III - 88% accuracy in predicting oncology trial failure and 89% accuracy in predicting trial success

Figure 4 below illustrates prediction accuracies, for both trial successes and failures, for five therapeutic areas. It should be noted that other therapeutic areas analysed show similar accuracy levels.

We would expect the intelligent forecasting outputs from the peak sales and probability of technical and regulatory success models to be used by customers in conjunction with predictions from other data sources, real world data, and in house bottom-up product forecasts, to provide additional insight and inform strategic business decisions.

Figure 5

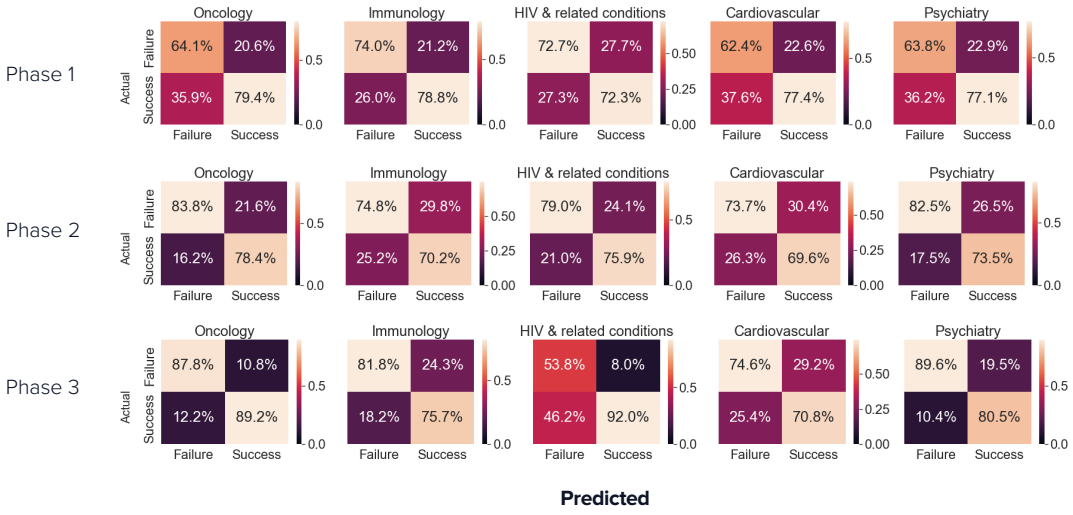

Figure 5: Confusion matrix by probability of technical and regulatory success for current NME. Boxes, and associated % accuracy measures, illustrate where intelligent forecasting using Evaluate's Omnium platform correctly predicted clinical trial success (bottom right) or trial failure (top left). Success rate is also illustrated by quadrant colour where dark shading is low accuracy and light shading is high accuracy.

In summary:

1. An intelligent forecasting approach deploying machine learning to analyse historical datasets, can identify signals of potential success for early, mid, and late-stage pipeline products and when combined with historical commercial datapoints, translate these insights into commercial value.

2. Peak sales predictions can be generated at individual product-level, dependent on therapeutic indication and market attributes. These product-level sales predictions are produced whether a development asset is in Phase I, II or III.

3. Peak sales outputs can also be produced at granular disease indication level.

4. Probability of technical and regulatory success predictions, i.e., clinical development success rates, can also be generated at product-level based on therapeutic indication. These product-level predictions are available for Phase I, II or III milestones at the start of each phase as well after phase transitions.

These observations illustrate how intelligent forecasting could provide advantages over other more conventional, established and widely used approaches, such as volume-, epidemiology- and consensus-based predictions. The benefits of intelligent forecasting include enhanced pipeline coverage, a balanced and fair approach to each individual product, and timely reactions and updates in response to news flow. Intelligent forecasting also allows products to be valued at scale and these estimates to be delivered quickly and accurately to inform time-sensitive strategic decision to be made.

Intelligent forecasting and the future – mirroring pharmaceutical companies' strategic direction to improving patient outcomes

Intelligent forecasting can go far beyond predicting probabilities of success and forecasting commercial opportunities for products in clinical development and providing competitive insights on disease areas or product pipelines of interest. Examples of where the application of intelligent forecasting will provide real world decision support across the life cycle of a therapeutic product are described below. These examples illustrate how this methodology can be used to overcome many of the historical and emerging forecasting challenges and address evolving market demand.

Here are four additional areas of pharmaceutical development where intelligent forecasting will add value:

1. **Preclinical/Early-stage R&D predictions** – In recent decades the focus for pharmaceutical companies has shifted to be further back in the development pipeline to try and identify successful candidates which have significant potential early on in their development cycle. These insights can be used to obtain first-mover advantages and to expedite the development of the most promising candidates. As a result, intelligence and insights on pre-clinical assets has become an increasingly useful and predictions of indication or therapeutic target success rates and/or target knowledge maps could be of significant value to life science organisations in the future. Applying machine learning to both historical and recent scientific publications, patent disclosures and scientific intelligence will be helpful in reducing both target and candidate attrition. Linking these foundational datasets with proprietary third-party data sets, such as pharmacovigilance and chemoinformatic databases will further add value to the generated insights.

2 **Predicting success, costs and timelines for innovative clinical trial designs** (and costs) both randomised clinical trials, pragmatic clinical trials, adaptive clinical trials, synthetic control arms in clinical trials, decentralized clinical trials, etc. Whilst drug development has always been a non-linear process, the number of approved products which take non-traditional routes to market is increasing as companies & regulators take steps to find efficiency savings across the drug development process to speed up access to life saving, or life changing, medicines for patients. This has resulted in the use of seamless study designs, decentralised trials, adaptive designs, and breakthrough therapy statuses to accelerate or skip phases of development. The changing nature of drug development will require that forecasting models are more flexible in the future.

3. **Taking intelligent forecasting to the next level by linking product/indication foundational data with epidemiology, market share, market access and product pricing data.** The inclusion of demand factors (e.g., diagnosis rate/treatment rate/willingness to pay) and/or epidemiology could be included to improve the outputs of the model as well as aligning the inputs to the expectations of clients. Expanding this to pricing, in particular net vs list pricing (users could import a discount) would also add a great deal of value for users. Precision medicine falls into this category i.e., predicting sales and success in clearly defined sub population.

4. **Real-world effectiveness.** There are differences between real world effectiveness of a medicine and the efficacy and safety endpoints that are collected during clinical trials. By building in news flow, regulatory information, medical publications and real-world evidence it will be possible to use intelligent forecasting to predict outcomes of medicines and understand how clinical trial endpoints translate into real world effectiveness e.g., predicting uplift from clinical trial efficacy to medical effectiveness and predicting outcomes. This will become increasing important as the life science and healthcare sectors move towards and adoption value-based, or outcome-based pricing for medicines.

Concluding comments

Commercial analytics and forecasting underpin strategic decision-making and business planning across the pharmaceutical and life science sectors. They are essential for managing and growing businesses, making informed decisions around future investments and portfolio prioritisation. Inaccurate risk assessments and poor strategic decision making are both costing the industry $ billions in wasted R&D spending every year and result in a gap in pharmaceutical intelligence to inform decision making in the clinical development space.

Intelligent forecasting can identify key risks and highlight return correlations that uncover insights into asset development and commercial opportunity by applying artificial intelligence, machine learning and predictive analytics to millions of data points curated from thousands of products across the entire clinical pipeline and commercial outcomes. Deploying intelligent forecasting to analyse historical datasets can identify signals of potential success, or failure, for early, mid, and late-stage pipeline products and when combined with historical commercial datapoints, translate these insights into commercial value and, therefore, accurately predict the future clinical and commercial outcomes.

Intelligent forecasting, in contrast to conventional and widely used forecasting methodologies, takes into account the important aspects of the pharmaceutical value chain such as the complexities of clinical development, the non-linear relationships between product attributes, the often randomness and uncertainties of drug discovery and market dynamics, as well as companies current development pipeline content and more importantly, they are adaptive and can be retrained to reflect recent changes more accurately, based on new evidence and data sets. Evaluate is one of the few providers in the market supplying real-time and granular product/ indication-level predictions for early- and mid- stage development assets.

Intelligent forecasting can provide real world decision support that customers can use to inform decision-making across the entire life cycle of a therapeutic product, including preclinical discovery, overcoming historical and emerging forecasting challenges.

Evaluate

a norstella company

Evaluate provides trusted commercial intelligence for the pharmaceutical industry. We help our clients to refine and transform their understanding of the past, present and future of the global pharmaceutical market to drive better decisions. When you partner with Evaluate, our constantly expanding solutions and our transparent methodologies and datasets are instantly at your disposal, along with personalised, expert support.

Evaluate gives you the time and confidence to turn understanding into insight, and insight into action.

Evaluate Pharma offers a global view of the pharmaceutical market's past, present and future performance with best-in-class consensus forecasts to 2028, unique broker forecasts, and the application of proprietary methodologies to support highly robust, detailed and accurate analysis.

Evaluate Omnium provides a complete, dynamic view of development risk and commercial return across all phases of the clinical lifecycle – including early-phase and privately-developed drugs not covered by analysts' forecasts. With product-specific data including Predicted Peak Sales, Probability of Technical and Regulatory Success (PTRS), R&D Costs, Net Present Value, Time-to-Peak and more, Evaluate Omnium makes it easier than ever to quantify and compare risk and return across the full pipeline landscape.

Evaluate Epi is curated by epidemiology experts and delivers comprehensive, global epidemiological data in granular detail, on a highly interrogatable platform. Customers have access to impartial data for 15 therapeutic areas, and over 230 indications and 9,500 sub-populations across 27 core markets (up to 49 for some countries).

Evaluate Medtech provides a transparent and trusted source of market intelligence and consensus forecasting for the global medical device and diagnostic landscape, using the same proprietary methodologies as Evaluate Pharma. Customers can quickly understand how the market views products and portfolios – and where their opportunities, risks and priorities lie.

Evaluate Consulting & Analytics are specialists in solving unique and complex biopharma pipeline, portfolio and commercialisation challenges with best-in-class datasets, powerful analytical capabilities, and deep therapy and commercialisation expertise.

Evaluate Vantage provides award-winning, thought-provoking news and insights on current and future developments in the pharma, biotech and medtech industries, and is the only news service underpinned by Evaluate's commercial intelligence and data.

www.evaluate.com | **@Evaluate** **@EvaluateVantage**

Evaluate Headquarters
Evaluate Ltd.
3 More London
London SE1 2RE
United Kingdom
T +44 (0)20 7377 0800

Evaluate Americas
EvaluatePharma USA Inc.
60 State Street, Suite 1910
Boston, MA 02109
USA
T +1 617 573 9450

Evaluate Asia Pacific
Evaluate Japan KK
Holland Hills Mori Tower 2F
5-11-2 Toranomon, Minato-ku
Tokyo 105-0001, Japan
T +81 (0)70 4131 0112